IN THE CLAIMS

This listing of the claim will replace all prior versions and listings of claim in the present application.

Listing of Claims

1. (currently amended)A similar document search method of searching ~~for~~ a ~~similar~~ document similar to a ~~specified~~ seeds document from a document database, the method, comprising the steps of:

a first extracting step of extracting at least one characteristic word ~~candidate as a candidate for a characteristic word~~ from ~~a~~ the seeds document including desired retrieval contents;

a second extracting step of extracting as characteristic words of the seeds document, ~~when~~ if the characteristic word ~~candidate~~ extracted by ~~said~~ the first extracting step is a compound characteristic word ~~including~~ constructed by a plurality of constituent characteristic words, the compound characteristic word and the constituent characteristic words included in the compound characteristic word ~~from the characteristic word candidate~~;

a step of calculating, according to the characteristic words extracted by ~~said~~ the second extracting step, similarity between the seeds document and ~~a registration document~~ the document stored on the document database, by using the characteristic words including the compound characteristic word and the constituent characteristic words by which the compound characteristic word is constructed; and

a step of outputting a retrieval result as a result of the similarity calculated by ~~said~~ the similarity calculating step.

2

2.   (currently amended)A similar document search method according to claim 1, wherein ~~said~~ the second extracting step includes a step of determining, when constituent characteristic word information indicating that constituent characteristic words ~~of~~ is the characteristic word is registered to the characteristic word ~~corresponding to the characteristic word candidate extracted by said first extracting step~~, that the characteristic word ~~candidate~~ is a compound characteristic word.

3.   (currently amended)A similar document search method according to claim 1, wherein ~~said~~ the similarity calculating step includes the steps of:

a step of calculating a weighting coefficient corresponding to a distance calculated by term appearance position on the seeds document between a constituent characteristic word and another constituent characteristic word which are extracted from one compound characteristic word; and

a step of calculating similarity by multiplying the weighting coefficient.

4.   (currently amended)A similar document search system for searching ~~for a similar~~ document similar to a ~~specified~~ seeds document from a document database, the system, comprising:

a document analyzer processor for extracting at least one characteristic word ~~candidate as a candidate for a characteristic word~~ from ~~a~~ the seeds document including desired retrieval contents;

a characteristic word extractor processor for extracting as characteristic words of the seeds document, ~~when~~ if the characteristic word ~~candidate~~ extracted by ~~said~~

3

the document analyzer processor is a compound characteristic word ~~including~~ constructed by a plurality of constituent characteristic words, the compound characteristic word and the constituent characteristic words included in the compound characteristic word ~~from the characteristic word candidate~~;

a seeds document similarity calculator processor for calculating, according to the characteristic words extracted by ~~said~~ the characteristic word extractor processor, similarity between the seeds document and ~~a registration document~~ the document stored on the document database, by using the characteristic words including the compound characteristic word and the constituent characteristic words by which the compound characteristic word is constructed; and

a retrieval result output processor for outputting a retrieval result as a result of the similarity calculated by ~~said~~ the seeds document similarity calculator processor.

5.   (currently amended)A similar document search system according to claim 4, wherein ~~said~~ the characteristic word extractor processor includes a compound characteristic word determiner processor for determining, when constituent characteristic word information indicating that constituent characteristic words ~~of~~ is the characteristic word is registered to the characteristic word ~~corresponding to the characteristic word candidate extracted by said document analyzer processor~~, that the characteristic word ~~candidate~~ is a compound characteristic word.

6.   (currently amended)A similar document search system according to claim 4, wherein ~~said~~ the seeds document similarity calculator processor includes:

4

a weighting coefficient calculator processor for calculating a weighting

coefficient corresponding to a distance <u>calculated by term appearance position on</u>

<u>the seeds document</u> between a constituent characteristic word and another

constituent characteristic word which are extracted from one compound

characteristic word; and

a calculator processor for calculating similarity by multiplying the weighting

coefficient.

7.   (currently amended)A program product for making a computer operate

as a similar document search system for searching ~~for a~~ ~~similar~~ document similar to

a ~~specified~~ <u>seeds</u> document <u>from a document database</u>, <u>the program</u> comprising:

a document analyzer processor program for extracting at least one

characteristic word ~~candidate as a candidate for a characteristic word~~ from ~~a~~ <u>the</u>

seeds document including desired retrieval contents;

a characteristic word extractor processor program for extracting as

characteristic words of the seeds document, ~~when~~ <u>if</u> the characteristic word

~~candidate~~ extracted by ~~said~~ <u>the</u> document analyzer processor program is a

compound characteristic word ~~including~~ <u>constructed by</u> a plurality of <u>constituent</u>

characteristic words, the compound characteristic word and <u>the</u> constituent

characteristic words included in the compound characteristic word ~~from the~~

~~characteristic word candidate~~;

a seeds document similarity calculator processor program for calculating,

according to the characteristic words extracted by ~~said~~ <u>the</u> characteristic word

extractor processor program, similarity between the seeds document and ~~a~~

5

~~registration document~~ the document stored on the document database, by using the characteristic words including the compound characteristic word and the constituent characteristic words by which the compound characteristic word is constructed; and

a retrieval result output processor program for outputting a retrieval result as a result of the similarity calculated by ~~said~~ the seeds document similarity calculator processor program.

8.   (currently amended)A program product for making a computer operate as a similar document search system according to claim 7, wherein ~~said~~ the characteristic word extractor processor program includes a compound characteristic word determiner processor program for determining, when constituent characteristic word information indicating that constituent characteristic words ~~of~~ is the characteristic word is registered to the characteristic word ~~corresponding to the characteristic word candidate extracted by said document analyzer processor program~~, that the characteristic word ~~candidate~~ is a compound characteristic word.

9.   (currently amended)A program product for making a computer operate as a similar document search system according to claim 7, wherein ~~said~~ the seeds document similarity calculator processor program includes:

a weighting coefficient calculator processor program for calculating a weighting coefficient corresponding to a distance calculated by term appearance position on the seeds document between a constituent characteristic word and another constituent characteristic word which are extracted from one compound characteristic word; and

6

a calculator processor program for calculating similarity by multiplying the

weighting coefficient.